RELATING REBELLION TO THE ENVIRONMENT:

AN ECONOMETRIC APPROACH

Edward J. Mitchell

November 1967
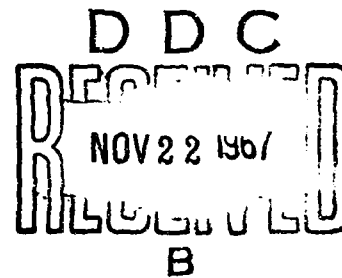
P-3726

9

RELATING REBELLION TO THE ENVIRONMENT:

AN ECONOMETRIC APPROACH

Edward J. Mitchell[*]

The RAND Corporation, Santa Monica, California

Knowledge of the conditions that give rise to rebellious behavior
should be among the most valuable assets of a policy-maker charged
with the security of an underdeveloped nation. Yet little in the way
of systematic research has been done relating rebellion to the circum-
stances that p duce and sustain it. This is not to say that there
aren't a number of existing hypotheses. Each time a policy decision
is made, there is presumably a corresponding theory being invoked
implicitly. This theory would typically be based upon experience,
introspection, or simple intuition.

What I would like to discuss here is a procedure by which one can
test theories about the determinants of rebellion. This approach
should be applicable to those nations that have experienced an insur-
gency and possess some data on the insurgency and the factors affecting
it. The approach has been used in Vietnam and is now being applied to
the post-war Huk uprising in the Philippines. From what I have been
able to discover about data sources, it should also be applicable to
Thailand.

One way of introducing this approach is to pose the following
questions: What are the observed characteristics of rebellious vil-
lages or provinces? What distinguishes these villages or provinces
from loyal ones? To each theory there corresponds some set of char-
acteristics. Testing the theory is largely a matter of comparing
hypothesized characteristics with actual ones. If one argues that
government neglect--that is, the relative absence of government
services--is one of the causes of rebellion, then villages under
rebel influence ought to be characterized by an absence of government
services. If poverty is supposed to contribute to rebellion, then
disloyal villages ought to be relatively poor. This must seem a
trivial and obvious requirement of a theory, yet this kind of test is
rarely if ever performed systematically and I suspect that many of the
commonly accepted notions about insurgency would not survive it.

One requirement is that we must be able to define and measure
"rebelliousness" or "disloyalty." In the Philippines the Constabulary
recorded various characteristics of barrios during the Huk insurgency.
I have, for example, a list of barrios in which the local officials
and the bulk of the peasantry actively supported (willingly or not)
the Hukbalahap. This is a fairly strict criterion. I hope to get a
list of barrios in which there was some degree of Huk influence. This
would presumably be a much larger list. As an overall measure of Huk
influence I have used thus far the percentage of barrios in a munici-
pality that met this strict criterion of Huk control. (A barrio
typically contains about 1000 people and there are about 20-25 barrios
per municipality. There are roughly 80 municipalities in Central
Luzon, the main area of Huk strength.) How in general one should
choose measures of rebel influence or government control is difficult
to say. This is really a question that goes well beyond the scope of
the present paper. In practical cases it is likely that this will
not present a serious problem of choice to the researcher, since the
army or police will probably have made it for him.

It is also clear that we must have measurements on those factors
that are supposed to promote rebellion. If poverty is thought to

contribute to insecurity, then we ought to have measures of real income:
for example, farm output per family, landlord-tenant shares, average
diets, etc. If certain government programs are thought to reduce the
likelihood of support for rebels, then data on variables such as miles of
roads, water systems, and educational opportunities should be available.
In no underdeveloped country will all the relevant data be at hand. The
Philippines is a rather good country for data. Thailand, from what I
have heard, is comparable to Vietnam, having had an agricultural but not
a population census. The significant results obtained in my Vietnam
study are, therefore, encouraging for a Thai study.

Given measurements on security and the factors that are theoreti-
cally supposed to determine security, we cannot simply compare or
correlate these measures. By itself a finding that higher literacy oc-
curs generally in secure areas does not imply that education has a favor-
able impact on security. First, it may be that literacy is merely
associated with some factor that does have a positive effect on security--
for example, income. The causation might run: higher incomes cause
greater security; higher incomes cause greater literacy. Literacy and
security may then be statistically correlated but there is no structural
relationship between them. Second, there may be a structural relation-
ship between literacy and security but the causation may run the other
way. Suppose that areas that had been insecure for some time did not
obtain teachers or educational equipment from the government and that
the rebels had poor educational programs. Literacy would be low in
insecure areas because security was low, not the other way round. Fi-
nally, there might be a pair of structural relationships between literacy
and security with causation running in both directions.[1]

These examples make it clear that simple correlations will often be
useless. The problem is that when there are many variables, and many
relationships among these variables, the full complexity of the world
cannot be captured in a simple correlation.

_____

[1]Incidentally, these examples have not been chosen for their
realism, but only for their pedagogical value.

The approach that I am suggesting is one that economists have been using for some time. The economic system is represented by a set of equations--behavioral, institutional, technological, and definitional. The variables in these equations are all those that the economist is interested in--prices, wages, employment--plus all those that must be taken into account if we are to explain the economic variables--family size, weather, technological conditions. A representative equation in an economic system would be the demand equation, or demand curve, relating the quantity of a good purchased to its price, the prices of other goods, income, family size, age, and so forth. Each decision made by "economic man" can be represented by a behavioral equation. The equation enables us to determine what an individual or group will choose to do, given the objective costs and benefits associated with the various choices and given the tastes or attitudes of those making the choices. In the demand equation the price of the good is an objective cost; the age of the consumer is a variable characterizing the consumer himself and therefore his likely tastes and attitudes.

I see no reason why the political system cannot be thought of in the same way. Whether a village is government- or rebel-controlled is a consequence of decisions made by people, individually or collectively, in that village. Their decision will be a function of the objective costs and benefits attached to alternative choices, and of the basic attitudes of these people toward the rebels and towards the existing order. There is then an equation relating the degree of security or control in an area to the "prices" of various actions in the area, and to those characteristics that determine the basic attitudes of the population.[1]

---

[1]Theories of insurgency probably differ not so much in stating how people will react to changes in costs or prices. We can generally agree that higher costs will tend to dissuade people from choosing a particular activity. Where they differ significantly is in positing what the characteristics of a rebel-held area will be. Is it likely to be poor, neglected by the government, with a generally uneducated population, or otherwise? This is the kind of question that we hope to be able to answer.

Unfortunately, estimation of the "rebellion equation" is not straightforward. There are variables affecting the disloyalty of a village that are also affected by disloyalty. The example given earlier shows that education can affect rebellion, and rebellion affect education, and also that both may affect each other simultaneously. There then exist additional structural relationships or equations between these variables. In this case the rebellion equation can be estimated properly only by taking into account these other relations. (In many cases the "rebellion equation" cannot be estimated at all due to the presence of these relations.) This is a very complex subject and I do not wish to go into it very far here. I shall instead focus on what can be done in a rather simple manner without requiring much in the way of a priori knowledge of all the equations involved.

It can be shown that it is appropriate to estimate a "rebellion equation" directly, providing that the variables determining rebellion may all be regarded as exogenous, that is, determined themselves by factors other than rebellion. Topographical variables are an excellent example. It may be argued that topography affects the likelihood of a village being rebellious. It is often said that guerrilla movements thrive in remote, inaccessible areas. But it is quite impossible to argue that rebellion affects topography. A region does not become mountainous because a rebellion is taking place there. Thus there can be no ambiguity about causation here.

Most variables assumed to be exogenous are not as obviously so as topography. Nevertheless, it will generally be possible to characterize variables as being primarily exogenous if any existing reverse causation is very weak.

If we can assume that this equation, with a measure of rebelliousness on one side and exogenous variables on the other, can be approximated by an equation linear in the parameters, then simple least-squares linear multiple regression can be applied to estimate the parameters and test hypotheses concerning them. To be concrete, let me give you some results from my Vietnam study. The variable to be explained there was the percentage of hamlets in a province under government control.

These hamlets were determined from a detailed map published in the
Los Angeles Times showing areas under government and Viet Cong control,
or contested. The map was derived from U.S. government sources. The
exogenous variables were obtained primarily from the 1960-61 Vietnamese
Census of Agriculture, and AID reports. They included measures of
tenancy, inequality in the sizes of land holdings, the presence of
large estates, land redistribution, population density, road density,
cross-country mobility, rice production per capita, and regional
effects.

The motivation for including land tenure variables in the equa-
tion was to examine the plausibility of the hypothesis that greater
inequalities in land tenure increase the likelihood of rebellious
behavior. This is certainly a very widely held expectation. The rice
production per capita variable is a proxy for real incomes and thus
helps us to evaluate the "poverty" hypothesis. All other variables
have fairly obvious theories associated with them.

After statistical estimation the equation, including only vari-
ables that were found to be statistically significant, looked like
this:

$$C = 6.47 - .36 \; OOL + 28.3 \; CV - 1.36 \; VL$$

$$+ .89 \; FL - .37 \; M + .09 \; PD,$$

$$\bar{R}^2 = .68,$$

where:

    C is the percentage of secure hamlets as calculated
from Los Angeles Times Map (1965)

    OOL is the percentage of all land that is owner-operated

    CV is the coefficient of variation of the distribution
of land holdings by size

    VL is the percentage of land subject to transfer under
the Diem Land Reform Program that was formerly
Vietnamese owned--approximately half of these estates
were redistributed by 1965

FL   is the percentage of land subject to transfer under
     the Diem Lan Reform Program that was formerly French
     owned--virtually none of these estates were redis-
     tributed by 1965

PD   is population density

M   is the percentage of area of good cross-country
     mobility

$\bar{R}^2$ is the coefficient of determination adjusted for
     degrees of freedom (an estimate of the proportion
     of the variance of control explained by the inde-
     pendent variables)

The principal conclusion to be drawn from the equation is that
inequality in land tenure implies greater government control. The
coefficients of each of the four land tenure variables suggest this.
Higher tenancy, more unequal distribution of farms by size and less
redistribution mean more control. Population density seems to con-
tribute positively to control whereas greater cross-country mobility
makes a negative contribution.

The six variables taken together explain about two-thirds of the
variance of control over twenty-six provinces. This explanation is
due primarily to the land tenure variables. Thus, not only do various
types of inequalities in land tenure appear to have an effect opposite
to that suggested by the hypothesis, but they in fact assume great
importance in this "perverse" role. This is not the place to discuss
in detail the interpretation of this phenomenon.[1] I would only say
that upon further reflection and reading of the historical literature
on rebellion the initial expectation is seen to have been unwarranted.
There are enough historical cases similar to the Vietnam situation to
indicate that the expectation should never have been confidently held.

There is no guarantee that the equation resulting from a study of
Thailand or any other country will be as provocative as the Vietnam

---

[1] The reader is referred to RM-5181-ARPA (Abridged), P-3610 (both
RAND publications), and my article in the August 1967 issue of _Asian
Survey_.

equation. Nevertheless, the Vietnam findings suggest that we may be entertaining notions about the relationship between rebellion and the environment that seriously disagree with reality. The only means by which we can improve our position is by further research. There are a number of recent rebellions and civil wars that could possibly be studied along the lines I have suggested in this paper. Besides Vietnam the list would include conflicts in the Philippines, China, and Thailand in Asia. The Spanish Civil War, and going back further into history, the English Civil War and French Revolution represent possible European studies. These researches may or may not produce much in the way of general theories about rebellions and revolutions, but again there is no way of determining this except by carrying them out.